Tomo-seq identifies SOX9 as a key regulator of cardiac fibrosis during ischemic injury

Lacraz, G.P.A., Junker, J.P., Gladka, M.M., Molenaar, B., Scholman, K.T., Vigil-Garcia, M., Versteeg, D., de Ruiter, H., Vermunt, M.W., Creyghton, M.P., Huibers, M.M.H., de Jonge, N., van Oudenaarden, A., van Rooij, E.

Circulation
2017 OCT 10 ; 136(15): 1396-1409
2017 JUL 19 (first published online: final publication)
doi: 10.1161/CIRCULATIONAHA.117.027832

Publisher: American Heart Association | Lippincott Williams & Wilkins

Tomo-seq identifies SOX9 as a key regulator of cardiac fibrosis during ischemic injury

Grégory P.A. Lacraz, MSc, PhD[1†]; Jan Philipp Junker, MSc, PhD[1,2†]; Monika M. Gladka, MSc, PhD[1]; Bas Molenaar, MSc[1]; Koen T. Scholman, MSc[1]; Marta Vigil-Garcia, MSc[1]; Danielle Versteeg, BS[1]; Hesther de Ruiter, BS[1]; Marit W. Vermunt, MSc, PhD[1]; Menno P. Creyghton, MSc, PhD[1] Manon M.H. Huibers, MSc, PhD[3]; Nicolaas de Jonge, MD[4]; Alexander van Oudenaarden, MSc, PhD[1] and Eva van Rooij, MSc, PhD[1,4*]

[1] Hubrecht Institute, KNAW and University Medical Center Utrecht, Utrecht, Netherlands
[2] Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany
[3] Dept. of Pathology. University Medical Center Utrecht, Utrecht, Netherlands
[4] Dept. of Cardiology. University Medical Center Utrecht, Utrecht, Netherlands

[†] These authors contributed equally to this work.

[*] Correspondence to:
E.vanrooij@hubrecht.eu
Hubrecht Institute, KNAW
Dept. of Cardiology, University Medical Center Utrecht
Uppsalalaan 8
3584CT Utrecht, Netherlands.
Tel: +31 30-2121800

Short title: Lacraz SOX9 as key regulator of cardiac fibrosis

Total word count: 4730

Abstract

Background Cardiac ischemic injury induces a pathological remodeling response, which can ultimately lead to heart failure. Detailed mechanistic insights into molecular signaling pathways relevant for different aspects of cardiac remodeling will support the identification of novel therapeutic targets.

Methods While genome-wide transcriptome analysis on diseased tissues has greatly advanced our understanding of the regulatory networks that drive pathological changes in the heart, this approach has been disadvantaged by the fact that the signals are derived from tissue homogenates. Here we used tomo-seq to obtain a genome-wide gene expression signature with high spatial resolution spanning from the infarcted area to the remote to identify new regulators of cardiac remodeling. Cardiac tissue samples from patients suffering from ischemic heart disease were used to validate our findings.

Results Tracing transcriptional differences with a high spatial resolution across the infarcted heart enabled us to identify gene clusters that share a comparable expression profile. The spatial distribution patterns indicated a separation of expressional changes for genes involved in specific aspects of cardiac remodeling, like fibrosis, cardiomyocyte hypertrophy, and calcium-handling (*Col1a2*, *Nppa*, and *Serca2*). Subsequent correlation analysis allowed for the identification of novel factors that share a comparable transcriptional regulation pattern across the infarcted tissue. The strong correlation between the expression levels of these known marker genes and the expression of the co-regulated genes could be confirmed in human ischemic cardiac tissue samples.

Follow-up analysis identified SOX9 as common transcriptional regulator of a large portion of the fibrosis-related genes that become activated under conditions of ischemic injury. Lineage-tracing

experiments indicated the majority of COL1-positive fibroblasts to stem from a pool of SOX9-expressing cells and *in vivo* loss of *Sox9* blunted the cardiac fibrotic response upon ischemic injury. The co-localization between SOX9 and COL1 could also be confirmed in patients suffering from ischemic heart disease.

Conclusions Based on the exact local expression cues, tomo-seq can serve to reveal novel genes and key transcription factors involved in specific aspects of cardiac remodeling. Using tomo-seq we were able to unveil the unknown relevance of SOX9 as key regulator of cardiac fibrosis, pointing to SOX9 as potential therapeutic target for cardiac fibrosis.

Keywords remodeling; ischemic heart disease; fibrosis; Sox9

Clinical Perspective

What Is New?

- SOX9 is a key regulator of cardiac fibrosis after ischemic injury in mice by regulating the expression of many extracellular matrix-related proteins.
- SOX9 is induced in cardiac tissue from patients suffering from ischemic heart disease and co-localizes with COL1 expression.
- Reduced levels of SOX9 lead to less cardiac fibrosis after ischemic injury in mice.
- Tomo-seq can be used to identify new players in cardiac biology and disease.

What Are the Clinical Implications?

- Our data suggest that therapeutic inhibition of SOX9 in the diseased heart could lead to a reduction in cardiac fibrosis.

Introduction

Ischemic heart disease induces a heterogeneous remodeling response across the damaged area that involves fibroblast activation, cardiomyocyte hypertrophy and changes in calcium handling, all of which are eventually detrimental for cardiac function.[1, 2] Fibroblast activation and cardiomyocyte hypertrophy occur as a direct effect of the local stress signals caused by the loss of viable tissue in the infarcted area. Subsequently, there is a decline in contractility of the surviving cardiomyocytes, which is caused by a change in metabolism and calcium handling genes.[3]

Genome-wide transcriptome analysis on extracts from diseased tissues has significantly enhanced our understanding of the gene regulatory networks that drive these pathological changes in the heart.[4, 5] However, to date, these approaches have been disadvantaged by the fact that the signals are derived from tissue homogenates, which inherently causes the loss of spatial information and dilutes out more localized expression signatures. Recent developments in RNA amplification strategies provide the opportunity to use small amounts of input RNA for genome-wide sequencing. Here we use tomo-seq[6] to obtain a genome-wide gene expression signature with high spatial resolution spanning from the infarcted area to the remote. Tracing transcriptional differences across the infarcted heart enabled us to identify clusters of genes with a comparable gene expression profile. In these individual clusters we recognized genes with well-known functions in specific aspects of heart remodeling, such as *Col1a2* for fibrosis, *Nppa* for cardiomyocyte hypertrophy, or *Serca2* for contractility. Correlation analyses using the spatial distribution patterns of these marker genes allowed for the identification of novel factors that share a comparable transcriptional regulation pattern across the infarcted tissue. Subsequent functional annotation analysis indicated that these genes could be linked to the known gene

function of their reference gene. The strong correlation between the expression levels of the markers genes and the expression of the co-regulated genes could be confirmed in human ischemic tissue samples.

Our data show that the high spatial resolution in gene expression signatures obtained by tomo-seq reveals new regulators, genetic pathways and transcription factors that are active in well-defined regions of the heart and potentially involved in specific aspects of heart disease. Using this technique, we identified SOX9 as a potent regulator of many of the *Col1a2* co-regulated genes. *In vivo* loss of *Sox9* reduced the expression of many extracellular matrix (ECM) genes which coincided with a blunted cardiac fibrotic response upon ischemic injury. These data unveil the currently unknown relevance of SOX9 as key regulator of cardiac fibrosis and underscores that tomo-seq can be used to increase our mechanistic insights into cardiac remodeling to help guide the identification of novel therapeutic candidates.

Methods

An expanded Methods section is available in the Supplemental Material online. Primers used to create ISH probes and for real-time PCR analysis are listed in the Supplemental Tables 7 and 8, respectively.

Ischemia reperfusion model

Animal experiments were performed in accordance with the institutional review committee at the Hubrecht Institute. Mice were randomly subjected to either sham or ischemia reperfusion surgery

as previously described.[7] Two weeks after surgery, cardiac tissue was collected for further analysis.

Tomo-seq

Tomo-seq experiments were performed as described elsewhere.[6] In short, 2.5 mm wide portions of cardiac mouse tissue spanning from the infarct towards the remote region of the left ventricular anterior wall were embedded in tissue freezing medium, frozen on dry ice, and cryosectioned into 48 slices of 80 μm thickness. We extracted RNA from individual slices and prepared barcoded Illumina sequencing libraries according to the CEL-seq protocol.[8] Paired-end reads obtained by Illumina sequencing were aligned to the transcriptome using BWA.[9] The 5' mate of each pair was used for mapping, discarding all reads that mapped equally well to multiple loci. The 3' mate was used for barcode information. Reads counts were normalized to the same number of total reads per section. Tomo-seq data analysis was performed in MATLAB (MathWorks) using custom-written code. For data analysis we used an expression cut-off of >4 reads in >1 section. In differential expression analysis (Figure 1C), we determined the boundary between remote and infarcted zone based on the spatial partitioning detected by pairwise comparison of sections across all genes in one biological replicate (Figure 1B). For the infarcted zone, we used sections 1-26, and for the remote zone we used sections 29-47. The border zone (sections 27-29) was omitted in order to reduce ambiguity in assignment of sections to zones. We then compared the sections within and outside the infarcted zone and assessed statistical significance with Wilcoxon rank sum test. For this analysis, each section was considered as an independent measurement. Furthermore, filtering was applied for genes that showed at least a two-fold expression difference between remote and infarcted zone. For this analysis, the mean

expression levels for each gene in the two zones was calculated. Concerning the hierarchical clustering, expression traces of the genes that passed the differential expression filter in Figure 1C were used for analysis. The data was standardized by *Z*-score normalization (along rows of data) so that the mean expression is zero and the standard deviation is 1 in order to remove differences in expression level between genes. Euclidean distance was used as distance metric. The assignment of genes to clusters I-III (Figure 1D) was determined manually considering the similarity in gene expression pattern across the ischemic heart.

SOX9 animal models

Sox9 (*Sox9*$^{fl/fl}$) mutant mice harboring two *loxP* sites flanking the exons 2-3 [10] were crossed with *Rosa26-CreERT2* mice (*R26*^{CreERT2}) to obtain an inducible *Sox9* loss-of-functional model (*Sox9*$^{fl/+}$*;R26*^{CreERT2}). For lineage tracing studies, mice expressing *CreERT2* under the control of the *Sox9* promoter [11] were bred with the *Rosa26-tdTomato* reporter mouse (*R26R*^{TdT}) to obtain *Sox9*^{CreERT2}*;R26R*^{TdT} mice. To induce the CreERT2 protein, *Sox9*$^{fl/+}$*;R26*^{CreERT2} and *Sox9*^{CreERT2};*R26R*^{TdT} mice were injected with Tamoxifen (corn oil/ethanol) intraperitoneally (2 mg at the day of surgery and 2 and 4 days after injury). Control mice (referred to as *Sox9*$^{fl/+}$*;R26*^{CreERT2} Vehicle) received an equal volume of the vehicle that was used to deliver Tamoxifen.

Pathway and transcription factor binding site enrichment

To investigate whether genes share a similar biological function, we searched for over-representation in the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway using

DAVID.[12] The enriched genes in the KEGG pathway are shown as *p* values corrected for multiple hypothesis testing using the Benjamini-Hochberg method.

Detection of over-represented conserved transcription factor binding sites in the set of genes spatially co-regulated to *Col1a2* was determined using single site analysis in oPOSSUM 3.0 (online software). The enrichment of SOX9 binding sites was determined using the *Z*-score, which uses the normal approximation to the binomial distribution to compare the rate of occurrence of a TFBS in the set of target genes to the expected rate estimated from the pre-computed background set.

Human heart samples

Approval for studies on human tissue samples was obtained from the Medical Ethics Committee of the University Medical Center Utrecht, The Netherlands (12#387). Written informed consent was obtained or in certain cases waived by the ethics committee when obtaining informed consent was not possible due to death of the patient. In this study, we included tissue from the left ventricular free wall of patients with end-stage heart failure secondary to ischemic heart disease. This end-stage heart failure tissue was obtained at explanation of the failing heart during heart transplantation or at autopsy. For each case, three areas of the infarcted heart tissue were included; 1) infarct zone, 2) border zone, and 3) remote area. For *in situ* hybridization (ISH) analysis, three patients were included. From these patients, the border zone of the infarcted hearts was used for ISH to verify tomo-seq. Gene expression values in infarct zone, border zone, and remote area obtained by real-time PCR were plotted for correlation analysis. Left ventricular free wall of non-failing donor hearts that could not be transplanted for technical reasons, were

used for comparison. In these cases, neither donor patient histories, nor echocardiography revealed signs of heart disease.

Statistical analysis

Values are presented as mean ± s.e.m. Previous studies were used to predetermine sample size. Statistical analyses between two groups were conducted using the two-tailed unpaired or paired Student's *t*-test or a Mann-Whitney test when the normality assumption was not met. Comparison among more than two groups was performed using a two-way ANOVA with Bonferroni's post-hoc test. Pearson's correlation coefficients were used to calculate gene pair correlation based on gene expression in human samples. KEGG pathways are ranked by their respective *p* value corrected for multiple hypothesis testing using the Benjamini-Hochberg method. *p* value <0.05 was interpreted to denote statistical significance. Prism 6 (GraphPad Software, Inc.) was used for statistical analyses.

Results

Tomo-seq performed on the infarcted mouse heart.

To obtain precise spatial information on local gene expression changes occurring in the heart in response to ischemic injury, we collected cardiac tissue from infarcted mice exposed to one hour of ischemia followed by either one or fourteen days of reperfusion (1 and 14 dpIR) and harvested tissue from sham-operated mice as control (Sham) (Figure 1A and Supplemental Figure 1).[7] Histological and molecular analysis confirmed a classical cardiac remodeling response in our model of ischemic injury, as exemplified by cardiac hypertrophy (Hematoxylin and Eosin staining, H&E), fibrosis (Sirius Red staining) and a change in expression of cardiac markers (Supplemental Figure 1).[7] Using microdissection, a small portion of the anterior wall of the left

ventricle spanning from the infarct towards the remote (2.5 mm wide and 4.0 mm long) was processed into ~50 consecutive cryosections with a thickness of 80 μm (Figure 1A). Subsequent RNA extraction from individual slices followed by RNA amplification, barcoding strategies and RNA sequencing [6] provided genome-wide data about the spatial distribution in gene expression across the injured heart (Supplemental Databases 1 through 3). A spatial partitioning between infarcted and remote area was visible at 1 and 14 dpIR, but not in the sham-operated samples when performing pairwise comparison of sections across all expressed genes (Figure 1B and Supplemental Figure 2A). The spatial separation became considerably more pronounced after filtering for genes that showed an at least two-fold and statistically significant differential expression between the infarct and remote zone by tomo-seq (Figure 1C and Supplemental Figure 2B). The number of regulated genes was found to be the highest 14 dpIR, which included 2357 coding genes and 134 non-coding transcripts (Figure 1D, Supplemental Figure 2C and Supplemental Databases 4 through 6). KEGG analysis on these regulated genes showed an enrichment for inflammatory pathway activation at day 1 after injury, while pathways involved in ECM, disease and cardiomyocyte remodeling were found to be regulated 14 dpIR (Supplemental Figure 2D and Supplemental Tables 1 and 2).

Gene expression patterns reveal localized remodeling responses.

Tracing transcriptional differences across the infarcted heart enabled us to identify clusters of genes with a comparable differential regulation throughout the infarcted heart at 14 dpIR (Figure 1D). The individual clusters contained well-known marker genes for specific aspects of heart remodeling, Collagen type I alpha 2 (*Col1a2*) (identified in cluster I), Natriuretic peptide A (*Nppa*) (identified in cluster II) and sarco/endoplasmic reticulum Ca^{2+}-ATPase (*Serca2*) (located

in cluster III). *Col1a2* is expressed in activated fibroblasts and important for cardiac fibrosis,[13] while *Nppa* is a cardiomyocyte-specific stress marker involved in myocyte hypertrophy.[14] Cardiomyocyte contractility is regulated by calcium fluxes to and from the sarcoplasmic reticulum and is impaired during heart disease. *Serca2* is a key regulator of Ca^{2+} transfer into the sarcoplasmic reticulum in muscle cells that is decreased during heart failure, which contributes to the decline in function.[3] The expression traces for *Col1a2*, *Nppa* and *Serca2* confirmed a gene-specific differential regulation from the infarcted area to the remote (Figure 1E). As expected *Col1a2* and *Nppa* were more abundantly expressed in the infarcted region 14 dpIR, while *Serca2* actually showed a decrease in expression towards the infarcted region (Figure 1E). ISH on cardiac tissue 14 dpIR confirmed the *Col1a2* expression to originate from activated fibroblasts, while the transcriptional peaks for *Nppa* stemmed from the stressed, hypertrophic cardiomyocytes immediately flanking the fibrotic regions (Figure 1F). We observed a decline in *Serca2* expression more towards the injured area, which is likely due to both a loss in cardiomyocytes as well as a decrease in transcriptional activation since the *Nppa* signals clearly indicate the presence of viable myocytes in this region (Figure 1F). The reproducibility of the obtained gene expression profiles was confirmed on a second set of samples (Supplemental Figure 3).

Tomo-seq identifies potential new players for cardiac remodeling and function.

An important advantage of tomo-seq over genome-wide sequencing techniques on tissue homogenates is that the local information on gene regulation allows for correlation analysis to identify genes with a comparable spatial distribution in transcriptional regulation.[6] Since we observed a gene-specific expression profile throughout the infarcted tissue for *Col1a2*, *Nppa* and

Serca2, we used the Euclidean distance of *Z*-score transformed spatial expression traces [6] to measure pattern similarity between genes 14 dpIR using *Col1a2, Nppa* and *Serca2* as reference genes. In doing so, we obtained a gene list that showed the greatest similarity in expressional differences across the infarcted tissue with our reference genes (Table 1, Figure 1G and Supplemental Databases 7 and 8), a vast majority of which could be identified within the corresponding gene cluster identified in Figure 1D. Interestingly, next to *Col1a2, Nppa* or *Serca2*, these lists also contained other well-known genes related to the biological function of the reference genes. Among the *Col1a2* co-expressed genes, we recognized additional genes known for their function in ECM deposition (like *Sparc* and *Col3a1*),[2] while many of the genes co-regulated with *Nppa* encode for proteins involved in cardiomyocyte hypertrophy (*Nppb* and *Myh7*). The gene list for *Serca*2 contained *Pln* and *Ryr2*, both well known for their importance in cardiac calcium handling and contractility [3] (Table 1, Supplemental Databases 7 and 8). This co-expression of genes could be confirmed by ISH and indicated the signals to stem from the same cell population (Figure 1H). The connection between the spatially co-regulated genes and their biological function was underscored by KEGG pathway analysis. The one hundred fifty genes with the highest similarity in expressional changes with either *Col1a2, Nppa* or *Serca2* throughout the infarcted heart at 14 dpIR, indicated an enrichment for the cellular function known to be associated with the reference genes (Figure 1I through K, Supplemental Tables 3 through 5, and Supplemental Databases 7 and 8). The known biological link of several of the listed genes and the functional connection based on gene ontology analysis suggests that the correlation analysis can serve to identify genes that are functionally related to the biological function of the reference genes.

RNA sequencing (RNA-seq) on whole tissue homogenates from the infarcted area from three

independent mice 14 dpIR showed a comparable directional regulation in gene expression, with the *Col1a2*- and *Nppa*-related genes going up after infarct, while the *Serca2*-related genes are going down compared to sham-operated mice (Supplemental Figure 4). However, in contrast with the data obtained by tomo-seq, the changes observed by RNA-seq on tissue homogenates failed to provide spatial information on co-expression of genes and showed smaller changes with a high inter-animal variation (Supplemental Figure 4). Since tomo-seq analysis is based on the correlation in gene expression within a single sample, the variation between animals is of lesser importance.

Tomo-seq analysis for lncRNAs specifically, showed localized expression changes, albeit far less pronounced and specific than for coding genes (Supplemental Figure 5), which is likely due to the low abundance of lncRNA transcripts.

The correlation in expression of novel genes linked to cardiac remodeling and function is conserved in humans.

While multiple well-known markers of fibrosis, hypertrophy and calcium handling could be identified among the genes with a similar transcriptional activation pattern, we also found multiple ill-studied genes that so far have not been linked to aspects of cardiac remodeling (Supplemental Databases 7 and 8). To confirm the correlation in transcriptional activation, we randomly chose one candidate from each list to explore in more detail. The *Z*-score transformed expression traces at 14 dpIR indicated a close correlation in expressional regulation between *Col1a2* with *Fstl1*, *Nppa* with *Pmepa1*, and *Serca2* with *Chchd2* (Figure 2A), which could be confirmed by ISH on murine cardiac tissue 14 dpIR (Figure 2B). Further confirmation for a

correlation in expression of these novel factors with *COL1A2*, *NPPA* and *SERCA2* was obtained by ISH on ischemic human heart tissue (Figure 2C).

The validity of using tomo-seq to identify genes that are expressionally linked was strengthened by the observation that real-time PCR analysis on cardiac tissue from patients suffering from ischemic heart disease confirmed the correlation between the expression levels of *COL1A2*, *NPPA, SERCA2* and the newly identified genes (Figure 2D through F). The correlation was strongly reduced when we cross-referenced genes from different lists (Supplemental Figure 6). The density plot for the cumulative Pearson correlation coefficients validates the shift towards a higher correlation between genes that belong to the same list (*co-regulated*) compared to the genes that were not shown to be co-regulated (*randomized*) by tomo-seq (Figure 2G).

While it remains to be determined which of the newly defined genes are relevant for cardiac remodeling, the functional link between the co-regulated genes and the fact that we can validate the co-regulation in both mice and human, implies that tomo-seq allows for the identification of novel genes that are potentially relevant for specific aspects of pathological remodeling of the infarcted heart.

Tomo-seq identifies SOX9 as key transcription factor for cardiac fibrosis.

The overlap in differential expression throughout the infarcted heart triggered us to explore whether a common transcription factor (TF) could be responsible for the synchrony in transcriptional regulation of the different gene clusters. Using an *in silico* approach, we searched for TFs (using oPOSSUM 3.0) that contain one or more predicted binding site(s) in the promoter regions of the top thirty *Col1a2* co-regulated genes (Table 1 and Supplemental Table 6). Among

these factors, we identified SOX9 as a potential candidate. SOX9 is a TF that has been recognized for its role in chondrocyte differentiation.[10]

While so far unstudied in the adult heart, previous work showed that SOX9 has a potent function in fibrosis.[15] Expression trace analysis for *Sox9* revealed a strong spatial correlation with *Col1a2* (Figure 3A). ISH indicated *Sox9* to be expressed in the same region of the infarcted area as *Col1a2*, although at a much lower level (Figure 3B). Real-time PCR on tissues from infarcted mouse heart further confirmed *Sox9* upregulation in the infarct zone (Figure 3C). Based on the predicted binding site(s) in the promoter regions of multiple *Col1a2* co-regulated genes, its proposed function in liver fibrosis and the overlap in transcriptional regulation with *Col1a2* in the infarcted heart, we decided to further pursue SOX9 in cardiac fibrosis. The induction in *Sox9* expression was only observed 14 dpIR and restricted to the infarcted area (Supplemental Figure 7A and B). Staining for both SOX9 and COL1 indicated SOX9 protein to be detectable in the same region as COL1 (Figure 3D).

To start exploring the fate of SOX9-expressing cells in the infarcted heart, we employed a lineage tracing approach using a TdTomato reporter mouse model driven by the promoter of the *Sox9* gene (*Sox9*^{CreERT2};*R26R*^{TdT}) (Figure 3E). Fluorescence-activated cell sorting (FACS) performed on single cells isolated from the left ventricle indicated a significant elevation of the SOX9-TdT^{+} cell population 14 dpIR compared to Sham (Figure 3F). Immunostaining clearly showed a co-localization between SOX9-TdT^{+} and COL1-expressing cells, which were surrounded by cardiomyocytes (marked by alpha-actinin-2, ACTN2) (Figure 3G and H). A similar overlap in expression was observed between SOX9-TdT^{+} and cells labelled with two other fibroblast markers; periostin and vimentin (Supplemental Figure 8A and B, respectively).[16,

[17] These data demonstrated that SOX9 is predominantly active in the fibroblast population that repopulates the infarcted area after injury.

To further explore whether SOX9 is involved in the transcriptional activation of the *Col1a2* co-regulated genes, we treated fibroblasts with TGFβ1 after we exposed them to either an siRNA against *Sox9* or a control. In addition to a strong repressive effect on *Sox9* expression, we also observed a significant repressive effect on 6 out of 15 potential SOX9 targets listed as *Col1a2* co-regulated genes, with a general downward trend for the remaining genes (Supplemental Figure 7C). This was also true for additional fibrosis-related genes (Supplemental Figure 7D), indicating a global function for SOX9 in fibroblast activation.

Real-time PCR analysis in tissue samples from human ischemic hearts showed a significant correlation between the levels of expression of *COL1A2* and *SOX9* (Figure 4A). In agreement with our mouse data, real-time PCR indicated an expressional increase in *SOX9* expression towards the infarcted area (Figure 4B), where the majority of the fibrosis is located. ISH showed that a sub-population of *COL1A2*-positive cells was also positive for *SOX9* in human ischemic hearts (Figure 4C). Immunostaining further confirmed that SOX9-expressing cells were also positive for COL1 (Figure 4D and E).

SOX9 regulates cardiac fibrosis during ischemia reperfusion injury.

To examine the effect of SOX9 *in vivo,* we generated inducible *Sox9* heterozygous knockout mice ($Sox9^{fl/+};R26^{CreERT2}$) (Figure 5A). Tamoxifen injection at the day of surgery and 2 and 4 days after injury resulted in a disruption of the *Sox9* allele as confirmed by PCR on genomic DNA (Figure 5B), and further quantified by real-time PCR and immunofluorescence after IR in the infarcted region (Figure 5C through E). *Sox9* loss of function was accompanied with a

profound reduction in fibrosis, which was quantified by the amount of Sirius Red staining in the infarcted region (Figure 5F and G). Periostin (PSTN), a protein marking activated fibroblasts,[17] was also reduced in the infarcted *Sox9*$^{fl/+}$*;R26*^{CreERT2} mice treated with Tamoxifen, further confirming the importance of SOX9 as a key driver for fibrosis in the ischemic heart (Figure 5H). Expression analysis for 15 randomly selected *Col1a2* co-expressed genes showed an increase in expression in response to ischemic injury. Loss of *Sox9* resulted in a significant reduction in expression for 13 out of 15 genes 14 dpIR compared to control animals (Figure 5I and Supplemental Database 7: highlighted in yellow).

High expression levels of SOX9 have previously been described in chondrocytes and publically available SOX9 ChIP-seq data in this cell type[18] showed that 15 out of the 30 genes that were co-expressed with *Col1a2* in our study have at least one of their predicted binding sites directly occupied by SOX9 (for instance: *Col1a2*, *Fn1*, *Lum* and *Vim*; Table 1 and Supplemental Figure 9). Importantly, these sites were found enriched for the histone mark H3K27ac in the adult mouse heart (ENCODE dataset), which further demonstrates that these regions are active and open for transcription factors like SOX9 *in vivo*. Altogether, these data demonstrate that SOX9 has the ability to occupy the promoter region of ECM-related genes and may actively regulate these genes in the heart.

Discussion

Here we applied tomo-seq to obtain a genome-wide gene expression profile with a high spatial resolution throughout the mammalian heart after ischemic injury. Cardiac ischemia reperfusion damage induces a heterogeneous remodeling response that involves several key processes, like cardiac fibrosis, cardiomyocyte hypertrophy, and a change in calcium handling within the heart

muscle cells.[2, 3, 13, 14] Localized expressional differences of well-known markers genes for these remodeling processes allowed us to uncover novel genes that showed a comparable transcriptional regulation and that are linked to specific aspects of cardiac remodeling. Using this data set, we identified SOX9 as a key transcriptional regulator of ECM-related genes and showed that *in vivo* loss of *Sox9* after myocardial infarction blunted the cardiac fibrotic response upon ischemic injury.

While RNA sequencing techniques on tissue samples have been instrumental in defining genes relevant for cardiac remodeling and repair,[4, 5] so far these approaches have been disadvantaged by the fact that the signals are derived from tissue homogenates, which inherently causes the loss of spatial information and dilutes out more localized expression signatures. Additionally, conventional methods for defining localized changes in genes expression, like ISH or immunohistochemistry, are limited to a defined set candidate genes and do not allow for genome-wide screening for novel relevant gene candidates. Recent developments in RNA amplification strategies provide the opportunity to use small amounts of input RNA for genome-wide sequencing, as exemplified by tomo-seq.[6] While recent studies showed this method to provide insightful data for both the developing and injured zebrafish heart,[6, 19] our study for the first time shows the relevance for the mammalian heart after ischemic injury. Especially the transcriptional differences, introduced by the localized heterogeneity in remodeling throughout an individual infarcted heart, appeared to be valuable for the identification of clusters of genes that showed a comparable regulation in expression. For this study we focussed on genes that showed an equivalent transcriptional regulation pattern across the infarcted tissue as well-known functions in fibrosis, cardiomyocyte hypertrophy or contractility (*Col1a2*, *Nppa* or *Serca2*).[2, 3, 13, 14] Based on subsequent functional annotation analysis, expressional confirmation in human

ischemic tissue samples and functional *in vitro* and *in vivo* assays, we conclude that the high spatial resolution in gene expression signatures obtained by tomo-seq allows for the identification of new relevant factors for specific aspects of heart disease. While we were preparing our manuscript, it was also reported that *Fstl1*, one of our top *Col1a2* co-regulated genes, is important for cardiac fibroblast activation,[20] which further underscores the relevance of our approach for identifying new players in specific cardiac remodeling responses.

Using our tomo-seq data, we identified SOX9 as common transcription factor able to regulate the expression of the majority of the *Col1a2* co-regulated genes. SOX9 is a transcription factor essential for chrondrogenesis via the activation of many ECM genes.[21] In the heart, SOX9 is highly expressed during development where it promotes epithelial-to-mesenchymal transition and ECM organization during heart valve development.[22, 23] In the adult heart, SOX9 has been shown to play a role in valve calcification.[24, 25] While SOX9 has been implicated in the fibrotic response of the liver,[15] so far it was unknown to play a role in cardiac fibrosis. We show that SOX9 is induced in response to ischemic injury and that *in vivo* loss of SOX9 after myocardial infarction blunts the cardiac fibrotic response upon damage, revealing a previously unknown function for SOX9 in cardiac fibrosis. In addition, we show that SOX9 is mainly active in the fibroblast population that repopulates the infarcted area after injury.

In our efficacy studies, we make use of a reduction in SOX9 levels instead of complete deletion, which is sufficient to cause an effect on cardiac fibrosis after injury. An equally profound phenotype in the heart has been reported by others upon heterozygous deletion of *Klf6* and *Rock1*, two other key regulators of fibrosis. [26, 27] This suggests that the molecular mechanism that drive cardiac fibrosis are sensitive to small perturbations in gene expression. Since therapeutical targeting of SOX9 would also moderately lower expression levels, we think this genetic model

gives a good representation of what would happen when using an inhibitor of SOX9 in the clinic as a therapy for cardiac fibrosis.

Here we show that the high spatial resolution in gene expression signatures obtained by tomo-seq reveals new regulators, genetic pathways, and transcription factors that are active in well-defined regions of the heart and potentially involved in specific aspects of heart disease. This knowledge increases our mechanistic insights into cardiac remodeling and function, and will help guide the identification of novel therapeutic candidates. However, the applicability of this approach is far greater than ischemic heart disease and the remodeling aspects we now focused on, and can also serve to identify new relevant factors for many different biological processes and disease states.

Acknowledgments

We thank Jeroen Korving, Reinier van der Linden, Stefan van der Elst and Harry Begthel for technical assistance.

Sources of Funding

All reported gene expression data are available in the Supplemental Materials. This work was supported by grants from the European Research Council (ERC AdG 294325 GeneNoiseControl and ERC CoG 615708 MICARUS) and a grant from the Leducq Foundation.

Disclosures

None

References

1. Tarone G, Balligand JL, Bauersachs J, Clerk A, De Windt L, Heymans S, Hilfiker-Kleiner D, Hirsch E, Iaccarino G, Knoll R, Leite-Moreira AF, Lourenco AP, Mayr M, Thum T and Tocchetti CG. Targeting myocardial remodelling to develop novel therapies for heart failure: a position paper from the Working Group on Myocardial Function of the European Society of Cardiology. *Eur J Heart Fail*. 2014;16:494-508.
2. van Berlo JH, Maillet M and Molkentin JD. Signaling effectors underlying pathologic growth and remodeling of the heart. *J Clin Invest*. 2013;123:37-45.
3. Lehnart SE, Maier LS and Hasenfuss G. Abnormalities of calcium metabolism and myocardial contractility depression in the failing heart. *Heart Fail Rev*. 2009;14:213-224.
4. Herrer I, Rosello-Lleti E, Rivera M, Molina-Navarro MM, Tarazon E, Ortega A, Martinez-Dolz L, Trivino JC, Lago F, Gonzalez-Juanatey JR, Bertomeu V, Montero JA and Portoles M. RNA-sequencing analysis reveals new alterations in cardiomyocyte cytoskeletal genes in patients with heart failure. *Lab Invest*. 2014;94:645-653.
5. Liu Y, Morley M, Brandimarto J, Hannenhalli S, Hu Y, Ashley EA, Tang WH, Moravec CS, Margulies KB, Cappola TP, Li M and consortium MA. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics*. 2015;105:83-89.
6. Junker JP, Noel ES, Guryev V, Peterson KA, Shah G, Huisken J, McMahon AP, Berezikov E, Bakkers J and van Oudenaarden A. Genome-wide RNA Tomography in the zebrafish embryo. *Cell*. 2014;159:662-675.
7. Hullinger TG, Montgomery RL, Seto AG, Dickinson BA, Semus HM, Lynch JM, Dalby CM, Robinson K, Stack C, Latimer PA, Hare JM, Olson EN and van Rooij E. Inhibition of miR-15 protects against cardiac ischemic injury. *Circ Res*. 2012;110:71-81.
8. Hashimshony T, Wagner F, Sher N and Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*. 2012;2:666-673.
9. Li H and Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589-595.
10. Akiyama H, Chaboissier MC, Martin JF, Schedl A and de Crombrugghe B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev*. 2002;16:2813-2828.
11. Kopp JL, Dubois CL, Schaffer AE, Hao E, Shih HP, Seymour PA, Ma J and Sander M. Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. *Development*. 2011;138:653-665.
12. Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44-57.
13. Li AH, Liu PP, Villarreal FJ and Garcia RA. Dynamic changes in myocardial matrix and relevance to disease: translational perspectives. *Circ Res*. 2014;114:916-927.
14. Sergeeva IA, Hooijkaas IB, Van Der Made I, Jong WM, Creemers EE and Christoffels VM. A transgenic mouse model for the simultaneous monitoring of ANF and BNP gene activity during heart development and disease. *Cardiovasc Res*. 2014;101:78-86.
15. Hanley KP, Oakley F, Sugden S, Wilson DI, Mann DA and Hanley NA. Ectopic SOX9 mediates extracellular matrix deposition characteristic of organ fibrosis. *J Biol Chem*. 2008;283:14063-14071.
16. Chang HY, Chi JT, Dudoit S, Bondre C, van de Rijn M, Botstein D and Brown PO. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc Natl Acad Sci U S A*. 2002;99:12877-12882.

17. Snider P, Standley KN, Wang J, Azhar M, Doetschman T and Conway SJ. Origin of cardiac fibroblasts and the role of periostin. *Circ Res*. 2009;105:934-947.
18. Ohba S, He X, Hojo H and McMahon AP. Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte. *Cell Rep*. 2015;12:229-243.
19. Wu CC, Kruse F, Vasudevarao MD, Junker JP, Zebrowski DC, Fischer K, Noel ES, Grun D, Berezikov E, Engel FB, van Oudenaarden A, Weidinger G and Bakkers J. Spatially Resolved Genome-wide Transcriptional Profiling Identifies BMP Signaling as Essential Regulator of Zebrafish Cardiomyocyte Regeneration. *Dev Cell*. 2016;36:36-49.
20. Maruyama S, Nakamura K, Papanicolaou KN, Sano S, Shimizu I, Asaumi Y, van den Hoff MJ, Ouchi N, Recchia FA and Walsh K. Follistatin-like 1 promotes cardiac fibroblast activation and protects the heart from rupture. *EMBO Mol Med*. 2016;8:949-966.
21. Ikeda T, Kamekura S, Mabuchi A, Kou I, Seki S, Takato T, Nakamura K, Kawaguchi H, Ikegawa S and Chung UI. The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. *Arthritis Rheum*. 2004;50:3561-3573.
22. Akiyama H, Chaboissier MC, Behringer RR, Rowitch DH, Schedl A, Epstein JA and de Crombrugghe B. Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. *Proc Natl Acad Sci U S A*. 2004;101:6502-6507.
23. Lincoln J, Kist R, Scherer G and Yutzey KE. Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. *Dev Biol*. 2007;305:120-132.
24. Peacock JD, Levay AK, Gillaspie DB, Tao G and Lincoln J. Reduced sox9 function promotes heart valve calcification phenotypes in vivo. *Circ Res*. 2010;106:712-719.
25. Huk DJ, Austin BF, Horne TE, Hinton RB, Ray WC, Heistad DD and Lincoln J. Valve Endothelial Cell-Derived Tgfbeta1 Signaling Promotes Nuclear Localization of Sox9 in Interstitial Cells Associated With Attenuated Calcification. *Arterioscler Thromb Vasc Biol*. 2016;36:328-338.
26. Rikitake Y, Oyama N, Wang CY, Noma K, Satoh M, Kim HH and Liao JK. Decreased perivascular fibrosis but not cardiac hypertrophy in ROCK1+/- haploinsufficient mice. *Circulation*. 2005;112:2959-2965.
27. Sawaki D, Hou L, Tomida S, Sun J, Zhan H, Aizawa K, Son BK, Kariya T, Takimoto E, Otsu K, Conway SJ, Manabe I, Komuro I, Friedman SL, Nagai R and Suzuki T. Modulation of cardiac fibrosis by Kruppel-like factor 6 through transcriptional control of thrombospondin 4 in cardiomyocytes. *Cardiovasc Res*. 2015;107:420-430.

Figure legends

Figure 1. High resolution gene expression atlas of the infarcted heart by tomo-seq. A, Schematic representation of a mouse heart after sham surgery (Sham) and 14 days post ischemia reperfusion (14 dpIR). **B**, Pairwise correlation for all sections across all genes showing clusters of correlated sections 14 dpIR in one biological replicate. **C**, Pairwise correlation for all sections across genes exhibiting at least two-fold and statistically significant differential expression between the infarct and remote zones 14 dpIR. **D**, Hierarchical clustering of expression traces for all genes that were found to be differentially expressed in **C**. **E**, Spatial expression pattern of three reference genes *Col1a2*, *Nppa*, and *Serca2* in the hearts from Sham, 1 dpIR, and 14 dpIR mice. **F**, Validation of the expression pattern by ISH 14 dpIR. Four chamber view (*left*) and higher magnification (*right*) are shown. Scale bars, 1 mm (*left*) and 200 μm (*right*). **G**, Spatial expression traces of ten co-regulated genes 14 dpIR. Reference genes are shown in red, and ten most similar genes are shown in grey. Black bold traces show other known markers involved in fibrosis, hypertrophy, and contractility (*Sparc*, *Nppb*, and *Pln*, respectively). **H**, Validation of the co-expression pattern of *Col1a2*/*Sparc*, *Nppa*/*Nppb*, and *Serca2*/*Pln* by ISH. Scale bars, 200 μm. **I** through **K**, KEGG analysis showing the enriched pathways the top one hundred fifty genes *Col1a2* (**I**), *Nppa* (**J**), and *Serca2* (**K**) co-regulated genes are involved in.

Figure 2. Identification of novel genes involved in remodeling and function of the ischemic heart. A, Spatial expression traces of three selected novel genes co-regulated with *Col1a2* (*Fstl1*), *Nppa* (*Pmepa1*), or *Serca2* (*Chchd2*) in mice 14 dpIR in one biological replicate. Expression traces were normalized by *Z*-score transformation. **B**, Validation of co-expression of *Col1a2/Fstl1*, *Nppa/Pmepa1*, and *Serca2/Chchd2* by ISH in mice 14 dpIR. Scale bars, 200 μm.

C, Validation of the co-expression pattern of *COL1A2/FSTL1*, *NPPA/PMEPA1*, and *SERCA2/CHCHD2* by ISH on human ischemic heart tissue. Scale bars, 100 μm. **D** through **F**, Real-time PCR analysis of genes that are spatially co-regulated in mice 14 dpIR (see Supplemental Database 7) using human cardiac tissue from ischemic heart disease patients. Control hearts and remote, border-zone and infarct zones from ischemic hearts are plotted. Data are presented as log 2 transformed values. Pearson correlation (*r*) and significance of co-regulated gene expression is shown (n=27–34; p<0.05 is considered significant). **G**, Kernel density plot of Pearson *r* values of the correlation in expression between the four corresponding *COL1A2*, *NPPA*, and *SERCA2* co-regulated genes (co-regulated; n=12) *vs* genes that are not co-regulated, i.e., genes cross-referenced from different lists (randomized; n=24) (see Supplemental Figure 6). Dotted lines depict the mean of the *r* values of all correlated and non-correlated genes.

Figure 3. Identification of *Sox9* as a key regulator of fibrosis-related genes. A, Spatial co-expression of *Col1a2* and *Sox9* determined by tomo-seq in the heart 14 dpIR in one biological replicate. **B**, Validation of the co-expression of *Sox9* and *Col1a2* in mice determined by ISH 14 dpIR. Scale bars, 100 μm. **C**, Real-time PCR analysis of *Sox9* expression in infarct (I) and remote (R) cardiac regions. Data are presented as fold change over sham-operated control hearts (n=5–6; *p<0.05 *vs* sham). **D**, Validation of the co-expression of SOX9 and COL1 in the infarct/border-zone in mice determined by co-immunostaining 14 dpIR. Nuclei were counterstained with DAPI. White arrows point to cells expressing SOX9 in their nuclei (purple). Scale bars, 50 μm. **E**, Schematic representation of the lineage tracing strategy of *Sox9* expressing cells. Reporter mice conditionally expressing TdTomato driven by the *Sox9* promotor ($Sox9^{CreERT2};R26R^{TdT}$) were subjected to sham surgery or IR, injected with Tamoxifen at day 0,

2, and 4 days post-surgery and analysed after 14 days. **F**, FACS quantification of cardiac SOX9-TdT^{+} cells in the hearts from Sham and 14 dpIR mice (n=3–4; *p<0.05 *vs* healthy sham control hearts). **G** and **H**, Co-immunostaining against TdTomato (TdT) and ACTN2 (**G**) or COL1 (**H**) in the hearts from sham-operated mice and 14 dpIR. White stars in the Merge field indicate SOX9-TdT-positive regions. Scale bars, 1 mm (4-chamber view) and 50 μm (higher magnification).

Figure 4. SOX9 is expressed in the fibrotic region in human cardiac tissue. A, Pearson correlation of *SOX9* and *COL1A2* expression determined by real-time PCR analysis on cardiac patient tissue (n=30). Data are presented as log 2 transformed values. **B**, Real-time PCR analysis of *SOX9* expression in infarct (I) and remote (R) cardiac regions of ischemic tissue samples. Data are presented as fold change over healthy control hearts (n=3–10; *p<0.05). **C**, Validation of the co-expression pattern of *SOX9/COL1A2* by ISH in human ischemic cardiac tissue. Scale bars, 100 μm (*left*) and 50 μm (*right*). **D** and **E**, Co-immunostaining against SOX9 and ACTN2 (**D**) or COL1 (**E**) in the hearts from control individuals or patients suffering from ischemic heart disease (IHD). White arrows point to cells expressing SOX9. Scale bars, 50 μm.

Figure 5. Loss of *Sox9* in mice protects against cardiac fibrosis. A, Schematic representation of the targeting strategy for conditional *Sox9* deletion. *Sox9*$^{fl/+}$;*R26*^{CreERT2} were subjected to sham surgery or IR, injected with Tamoxifen at day 0, 2, and 4 days post-surgery and analysed after 14 days. **B**, PCR genotyping for *Sox9* floxed deleted allele (*Sox9fl del*) and *Cre* transgene. Genomic DNA isolated from *Sox9*$^{fl/+}$;*R26*^{CreERT2} treated with vehicle or Tamoxifen. Used forward (P1) and reverse primers (P2) are indicated as demi-arrowheads in **A**. M, Marker. **C**, Real-time PCR analysis of *Sox9* in the hearts (infarct zone) from *Sox9*$^{fl/+}$;*R26*^{CreERT2} mice 14

dpIR injected with either vehicle or Tamoxifen (n=4–7; *p<0.05 *vs* the indicated groups). **D**, Co-immunostaining against SOX9 and ACTN2 on corresponding infarcted heart tissue 14 dpIR. Right panel shows the same section including bright field (BF). Scale bars, 50 μm. **E**, Quantification of SOX9 expression 14dpIR in the fibrotic region of the left ventricle or the corresponding region in Sham mice (n=3; **p<0.01, *p<0.05 *vs* Sham). **F**, Histological sections of infarcted hearts stained for Sirius Red (collagen) 14 dpIR. Scale bars, 1 mm (*left*) and 400 μm (*right*). **G**, Quantification of Sirius Red-positive area in infarcted area or corresponding Sham region (n=3–7; **p<0.01; *p<0.05 *vs* corresponding Sham). **H**, Western blot analysis of the fibrotic protein periostin (PSTN) in the hearts from $Sox9^{fl/+};R26^{CreERT2}$ mice 14 dpIR. GAPDH was used as a loading control (n=4–5; *p<0.05 *vs* MI vehicle injected mice). **I**, Real-time PCR analysis of *Col1a2* co-expressed genes in the mouse hearts 14 dpIR (n=6; **p<0.01; *p<0.05 *vs* Sham injected with Vehicle).

Table 1. Top thirty genes showing the most similar expression pattern to *Col1a2*.

Col1a2 similar genes	SOX9 predicted binding site	SOX9 validated binding site
Col1a2	x	x
Sparc	x	
Fstl1	x	x
Serping1	x	
Pdgfrl	x	
Tmem45a	x	
Col3a1	x	x
Sfrp1	x	x
Lox	x	x
Ecrg4		
Dkk3	x	
Col1a1	x	
Itgbl1	x	
Fn1	x	x
Thbs2		
Cthrc1	x	
Col8a1	x	
Col5a2	x	
Lum	x	x
Fbln2	x	
Gas1	x	x
Antxr1	x	x
Thbs1	x	x
Ogn	x	
Col16a1	x	x
Vim	x	x
Cxcl16	x	
Timp1	x	x
Rnase4	x	x
Ddah1	x	x

Genes that contain a predicted/validated SOX9 binding site in their promoter region are marked.

Figure 1

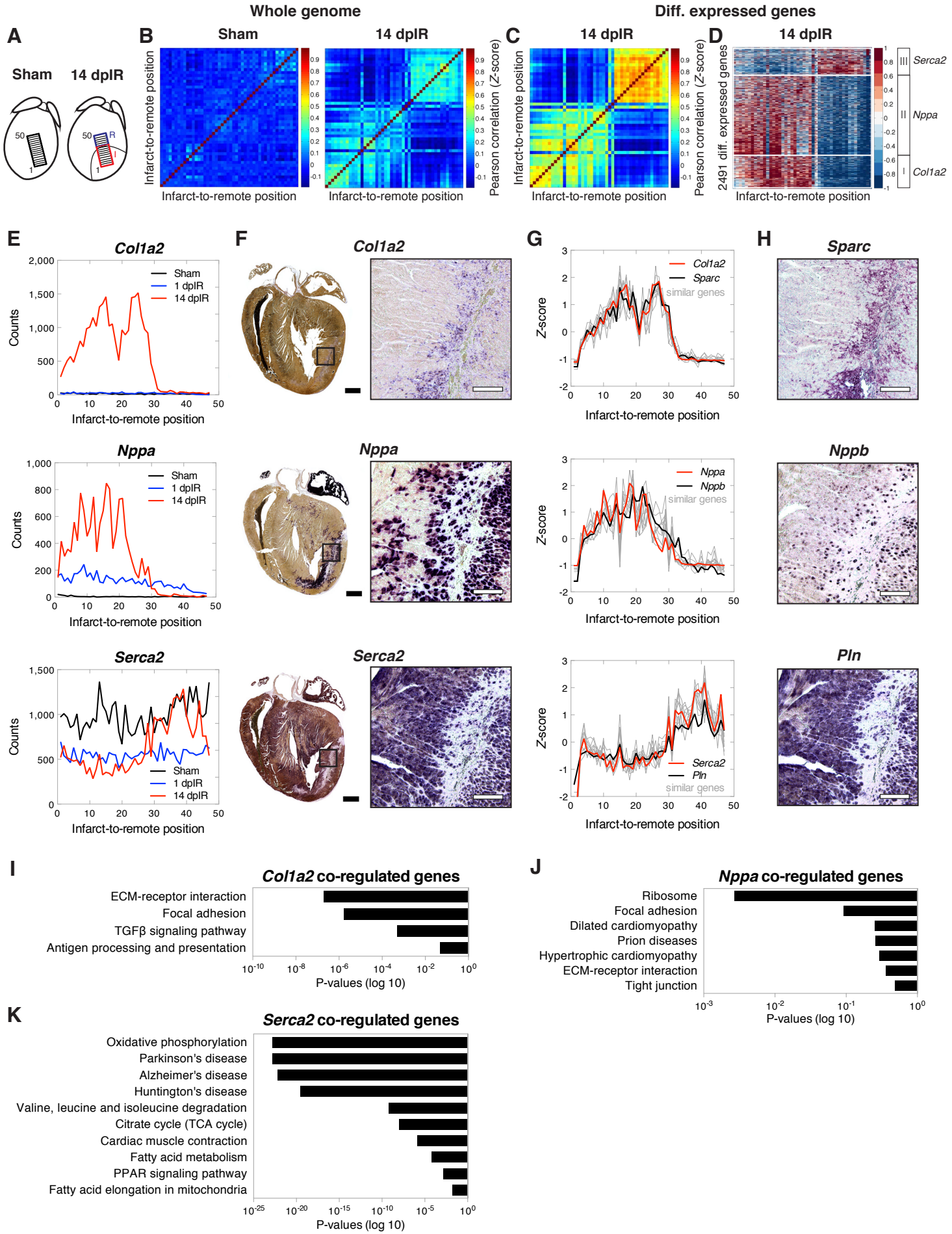

Whole genome
Diff. expressed genes
A
Sham
14 dpIR
B
Sham
14 dpIR
Infarct-to-remote position
Pearson correlation (Z-score)
C
14 dpIR
D
14 dpIR
2491 diff. expressed genes
III Serca2
II Nppa
I Col1a2
E
Col1a2
Counts
Sham
1 dpIR
14 dpIR
F
Col1a2
G
Col1a2
Sparc
similar genes
Z-score
H
Sparc
Nppa
Nppb
Serca2
Pln
I
Col1a2 co-regulated genes
ECM-receptor interaction
Focal adhesion
TGFβ signaling pathway
Antigen processing and presentation
P-values (log 10)
J
Nppa co-regulated genes
Ribosome
Focal adhesion
Dilated cardiomyopathy
Prion diseases
Hypertrophic cardiomyopathy
ECM-receptor interaction
Tight junction
K
Serca2 co-regulated genes
Oxidative phosphorylation
Parkinson's disease
Alzheimer's disease
Huntington's disease
Valine, leucine and isoleucine degradation
Citrate cycle (TCA cycle)
Cardiac muscle contraction
Fatty acid metabolism
PPAR signaling pathway
Fatty acid elongation in mitochondria

Figure 2

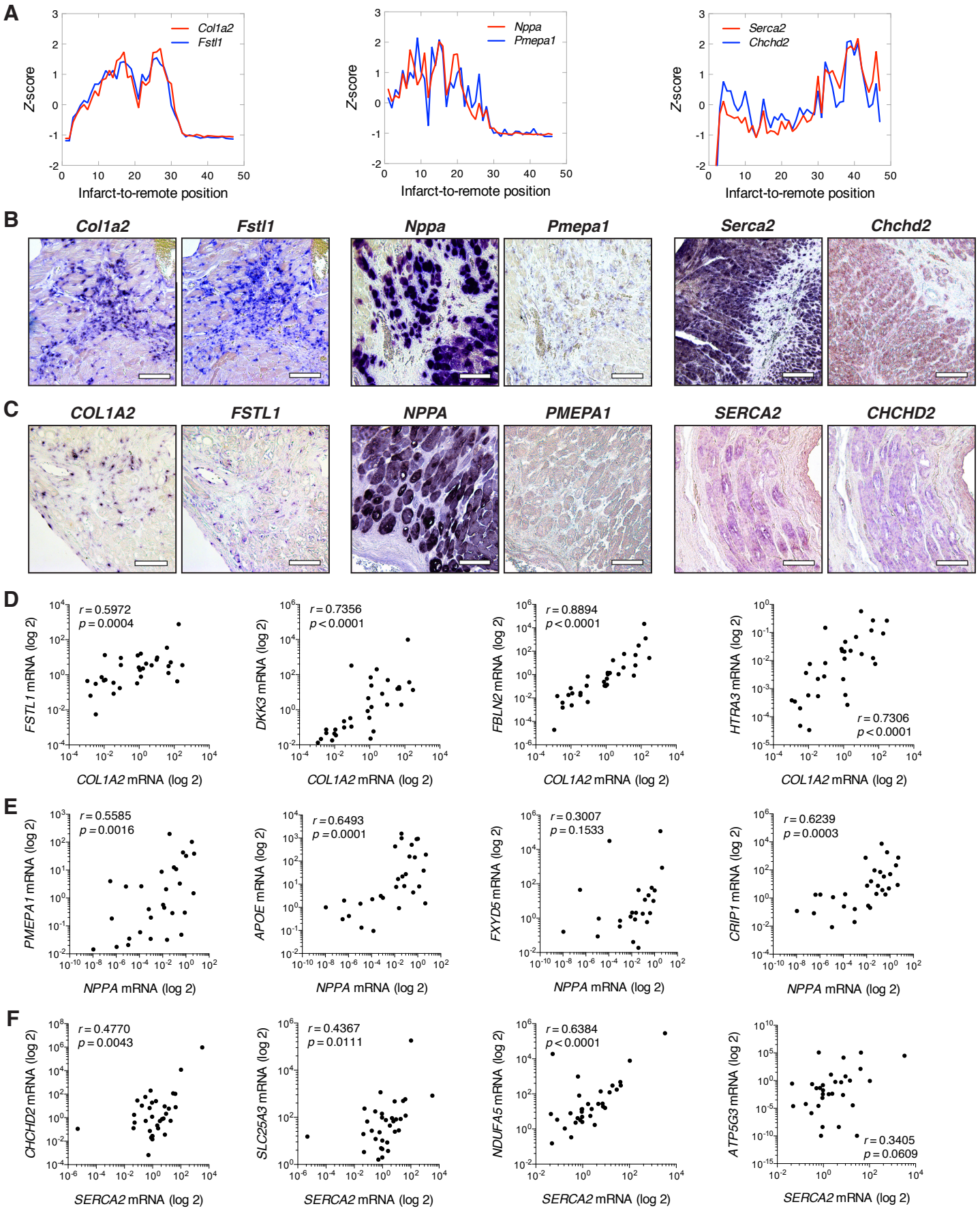

A
Col1a2
Fstl1
Nppa
Pmepa1
Serca2
Chchd2
Z-score
Infarct-to-remote position
B
Col1a2
Fstl1
Nppa
Pmepa1
Serca2
Chchd2
C
COL1A2
FSTL1
NPPA
PMEPA1
SERCA2
CHCHD2
D
FSTL1 mRNA (log 2)
DKK3 mRNA (log 2)
FBLN2 mRNA (log 2)
HTRA3 mRNA (log 2)
COL1A2 mRNA (log 2)
r = 0.5972
p = 0.0004
r = 0.7356
p < 0.0001
r = 0.8894
p < 0.0001
r = 0.7306
p < 0.0001
E
PMEPA1 mRNA (log 2)
APOE mRNA (log 2)
FXYD5 mRNA (log 2)
CRIP1 mRNA (log 2)
NPPA mRNA (log 2)
r = 0.5585
p = 0.0016
r = 0.6493
p = 0.0001
r = 0.3007
p = 0.1533
r = 0.6239
p = 0.0003
F
CHCHD2 mRNA (log 2)
SLC25A3 mRNA (log 2)
NDUFA5 mRNA (log 2)
ATP5G3 mRNA (log 2)
SERCA2 mRNA (log 2)
r = 0.4770
p = 0.0043
r = 0.4367
p = 0.0111
r = 0.6384
p < 0.0001
r = 0.3405
p = 0.0609
G
Density
r Coefficients
Randomized
Co-regulated

Figure 3

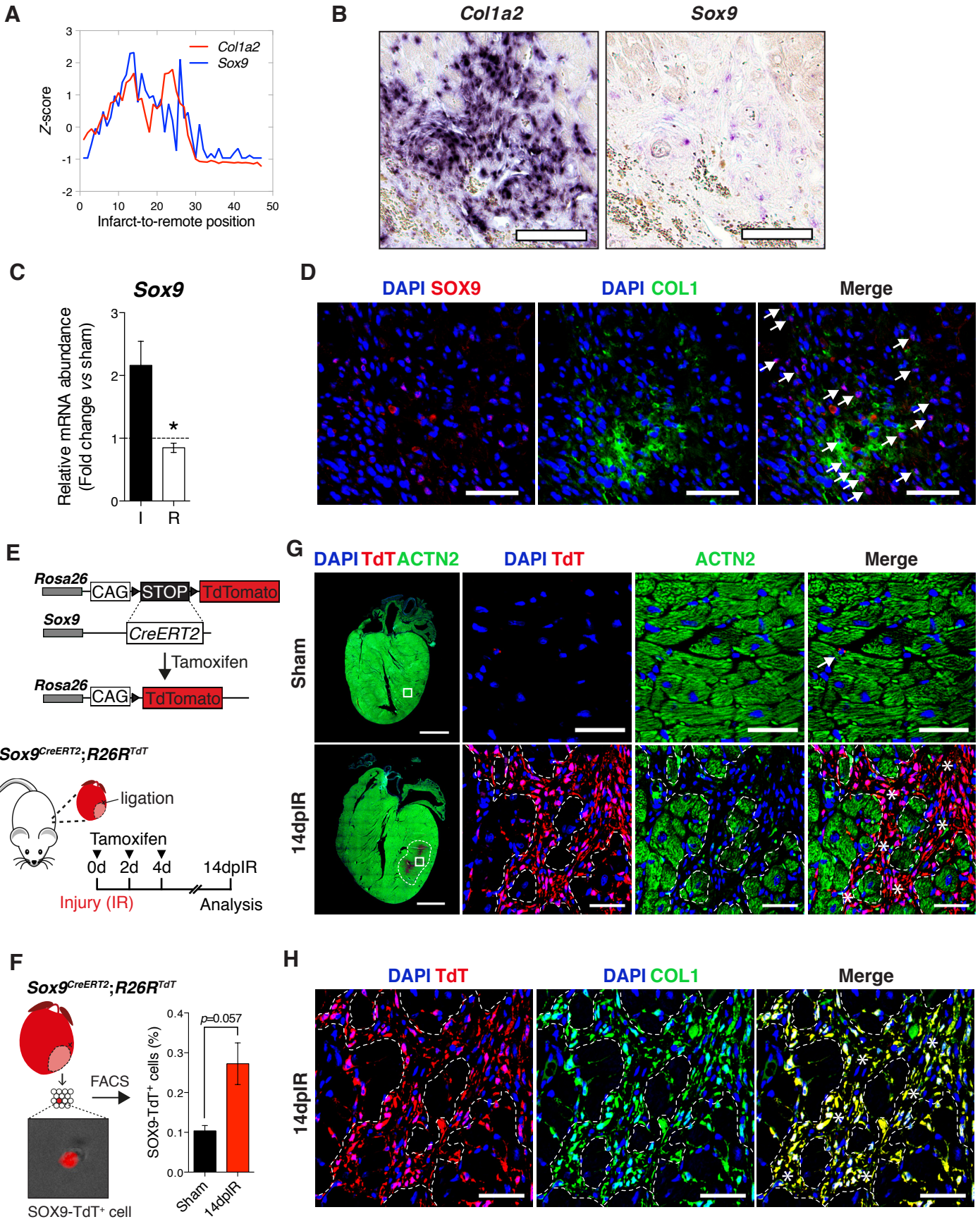

A
Col1a2
Sox9
Z-score
Infarct-to-remote position
B
Col1a2
Sox9
C
Sox9
Relative mRNA abundance
(Fold change vs sham)
I
R
D
DAPI SOX9
DAPI COL1
Merge
E
Rosa26
CAG
STOP
TdTomato
Sox9
CreERT2
Tamoxifen
Rosa26
CAG
TdTomato
Sox9CreERT2;R26RTdT
ligation
Tamoxifen
0d
2d
4d
14dpIR
Injury (IR)
Analysis
G
DAPI TdT ACTN2
DAPI TdT
ACTN2
Merge
Sham
14dpIR
F
Sox9CreERT2;R26RTdT
FACS
SOX9-TdT+ cell
SOX9-TdT+ cells (%)
p=0.057
Sham
14dpIR
H
DAPI TdT
DAPI COL1
Merge
14dpIR

Figure 4

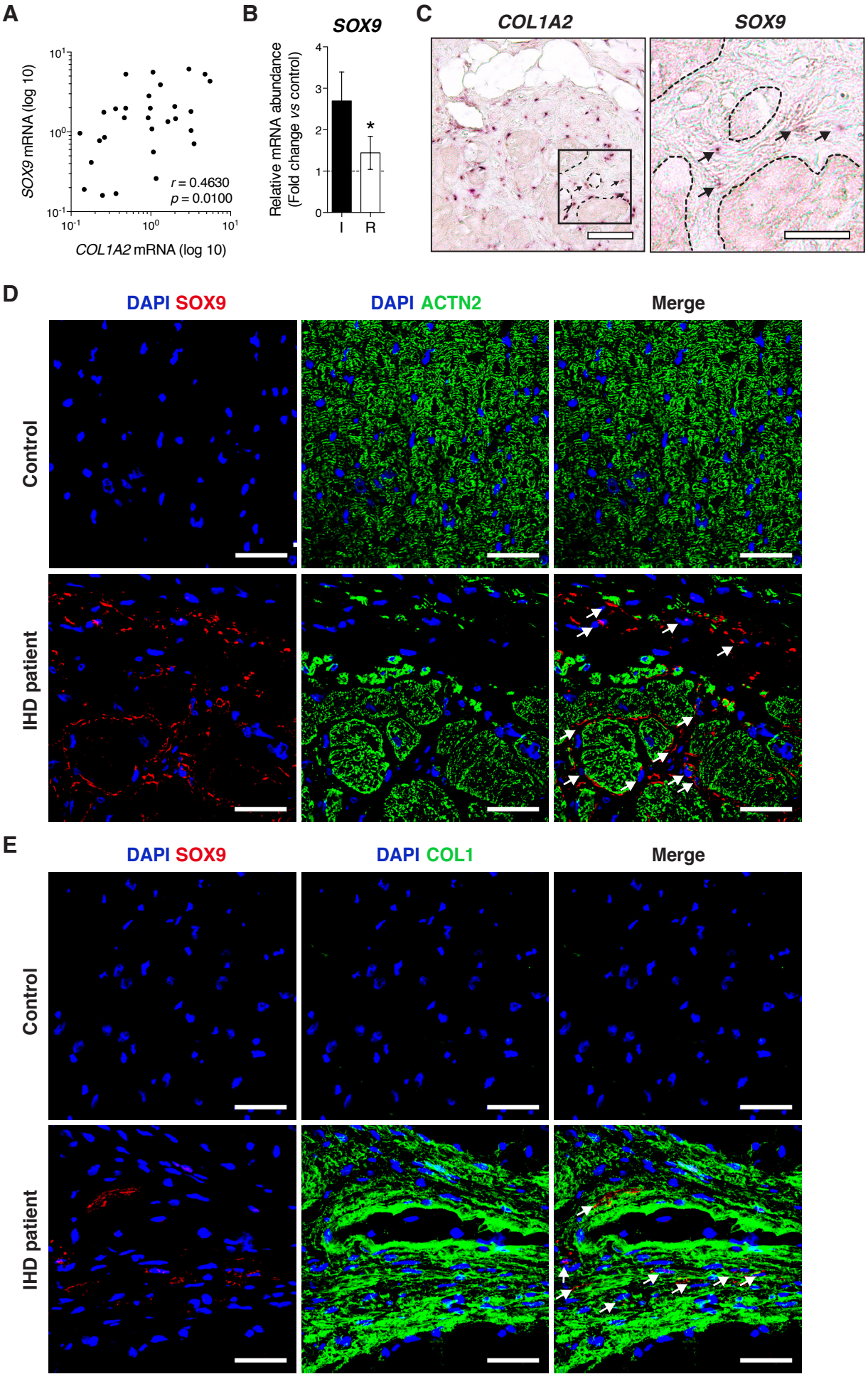
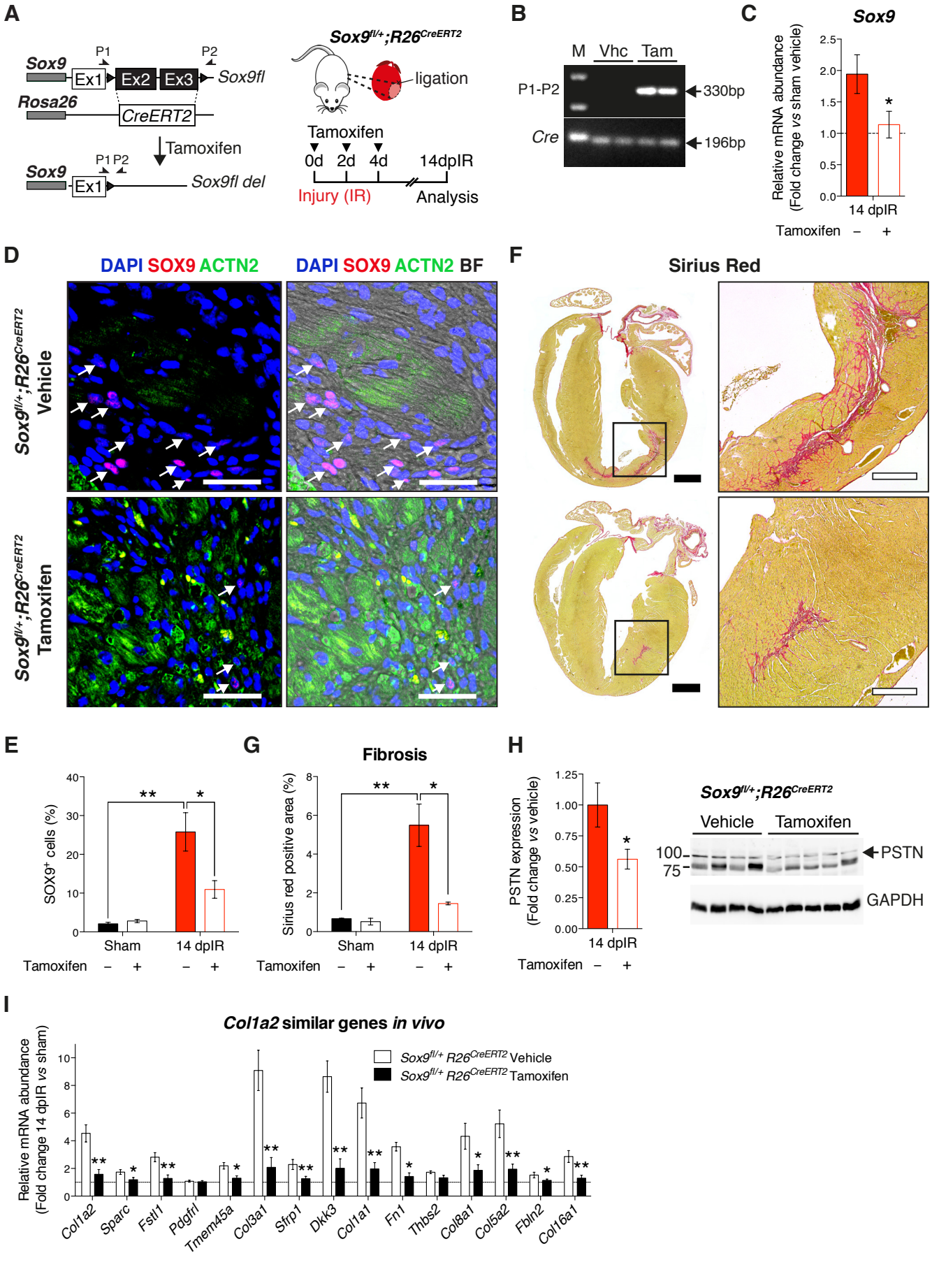

A
SOX9 mRNA (log 10)
COL1A2 mRNA (log 10)
r = 0.4630
p = 0.0100
B
SOX9
Relative mRNA abundance
(Fold change vs control)
I
R
C
COL1A2
SOX9
D
DAPI SOX9
DAPI ACTN2
Merge
Control
IHD patient
E
DAPI SOX9
DAPI COL1
Merge
Control
IHD patient

Figure 5


A
Sox9
P1
P2
Ex1
Ex2
Ex3
Sox9fl
Rosa26
CreERT2
Tamoxifen
Sox9
P1 P2
Ex1
Sox9fl del
Sox9fl/+;R26CreERT2
ligation
Tamoxifen
0d 2d 4d
14dpIR
Injury (IR)
Analysis
B
M Vhc Tam
P1-P2
330bp
Cre
196bp
C
Sox9
Relative mRNA abundance
(Fold change vs sham vehicle)
14 dpIR
Tamoxifen – +
D
DAPI SOX9 ACTN2
DAPI SOX9 ACTN2 BF
Sox9fl/+;R26CreERT2
Vehicle
Sox9fl/+;R26CreERT2
Tamoxifen
F
Sirius Red
E
SOX9+ cells (%)
Sham
14 dpIR
Tamoxifen – + – +
G
Fibrosis
Sirius red positive area (%)
Sham
14 dpIR
Tamoxifen – + – +
H
PSTN expression
(Fold change vs vehicle)
14 dpIR
Tamoxifen – +
Sox9fl/+;R26CreERT2
Vehicle
Tamoxifen
100
75
PSTN
GAPDH
I
Col1a2 similar genes in vivo
Relative mRNA abundance
(Fold change 14 dpIR vs sham)
Sox9fl/+ R26CreERT2 Vehicle
Sox9fl/+ R26CreERT2 Tamoxifen
Col1a2
Sparc
Fstl1
Pdgfrl
Tmem45a
Col3a1
Sfrp1
Dkk3
Col1a1
Fn1
Thbs2
Col8a1
Col5a2
Fbln2
Col16a1